

Report by: Anand Prashar

NLP Homework 5 – q1

Q:

Please complete `cossim_sparse(v1,v2)` in `distsim.py` to compute and display the cosine similarities between each pair of these words. Briefly comment on whether the relative similarities make sense.

A:

Method has been implemented and uploaded on Vocareum.

Below is the result of executing provided file: `testq1.py`

```
<terminated> testq1.py [C:\Anaconda\python.exe]
file nytcounts.university_cat_dog has contexts for 3 words
Cosine similarity between cat and dog 0.966891672715
Cosine similarity between cat and university 0.660442421144
Cosine similarity between university and dog 0.659230248969
```

Comment:

Cat and Dog share high cosine similarity here;

University shares a lower similar similiarity(~ 0.66) with both Cat and Dog.

It makes sense as the vectors used were *context counts*, and we hear Cat and Dog in same context much often than hearing them in context with University.